



Derived Variables

Revised by FHWA, August 2020

Submitted to:
Federal Highway Administration
Office of Policy Information
1200 New Jersey Avenue, SE
Washington, DC 20590

Contract # GS23F8144H
Order # DTFH6114F00113

Submitted by:
Westat
1600 Research Boulevard
Rockville, Maryland 20850-3129
(301) 251-1500

HOUSEHOLD Variables

CDIVMSAR

Description: Grouping of household by combination of census division, MSA status, and presence of a subway system when population greater than 1 million

Logic: CENSUS_D & MSACAT

CENSUS_D

Description: 2010 Census division classification for the respondent's home address

Logic: [DIVISIONS] = http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

```
IF HHSTATE = ("ME","NH","VT","CT","MA","RI") THEN 1
IF HHSTATE = ("NY","NJ","PA") THEN 2
IF HHSTATE = ("IL","IN","MI","OH","WI") THEN 3
IF HHSTATE = ("IA","KS","MO","MN","ND","NE","SD") THEN 4
IF HHSTATE = ("DC","DE","FL","GA","MD","NC","SC","WV","VA") THEN 5
IF HHSTATE = ("AL","KY","MS","TN") THEN 6
IF HHSTATE = ("AR","LA","OK","TX") THEN 7
IF HHSTATE = ("AZ","CO","ID","MT","NM","NV","UT","WY") THEN 8
IF HHSTATE = ("AK","CA","HI","OR","WA") THEN 9
```

Note: Census_D is formulated based on states of original sampling.

CENSUS_R

Description: Census region classification for home address

Logic: [REGIONS] = http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

```
IF CENSUS_D = (1,2) THEN 1
IF CENSUS_D = (3,4) THEN 2
IF CENSUS_D = (5,6,7) THEN 3
IF CENSUS_D = (8,9) THEN 4
```

CNTTDHH

Description: Count of household trips on travel day

Logic: COUNT(TDTRPNUM)

DRVRCNT

Description: Number of drivers in household

Logic: COUNT(WHERE DRIVER = 1)

HH_CBSA

Description: Core Based Statistical Area (CBSA) FIPS code for the respondent's home address

Logic: CBSA based on the household's home geocode and TIGER/Line® Shapefile

geometries.

Source: [https://www.census.gov/cgi-](https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2014&layergroup=Core+Based+Statistical+Areas)

[bin/geo/shapefiles/index.php?year=2014&layergroup=Core+Based+Statistical+Areas](https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2014&layergroup=Core+Based+Statistical+Areas)

HH_HISP

Description: Hispanic status of household respondent

Logic: R_HISP WHERE PERSONID = 1

HH_RACE

Description: Race of household respondent

Logic: R_RACE WHERE PERSONID = 1

HHRELATD

Description: At least two household persons are related

Logic: IF COUNT(R_RELAT) WHERE R_RELAT = (2,3,4,5,6) > 0 THEN 1
ELSE 2

HHRESP

Description: Person identifier of household respondent

Logic: SELECTPERSON1

HHSIZE

Description: Count of household members

Logic: COUNT(PERSONID)

HHSTATE

Description: Household state

Logic: STATE WHERE LOCATION IS HOME

HHSTFIPS

Description: State FIPS for household address

Logic: STATEFIPS WHERE LOCATION IS HOME

HHVEHCNT

Description: Count of household vehicles

Logic: COUNT(VEHID WHERE VEHTYPE = 1,2,3,4,5,6,7)

LIF_CYC

Description: Life Cycle classification for the household, derived by attributes pertaining to age, relationship, and work status.

Logic: IMPUTED_AGE =
IF AGE IS MISSING OR AGE = (-7,-8) THEN
IF AAGE IS MISSING OR AAGE = (-7,-8) THEN 41

```

IF AAGE = 1 THEN 2
IF AAGE = 2 THEN 10
IF AAGE = 3 THEN 16
IF AAGE = 4 THEN 41
IF AAGE = 5 THEN 70
IF AAGE = 6 THEN 77
ELSE AGE

ADULT_CHILD_STATUS =
IF IMPUTED_AGE < 18 THEN "CHILD"
IF IMPUTED_AGE > 21 THEN "ADULT"
IF IMPUTED_AGE BETWEEN 18 AND 21 THEN
IF R_RELAT = 3 THEN "CHILD"
IF R_RELAT = (1,5,6) AND [ANY OTHER HH MEMBER] R_RELAT = 4 THEN "CHILD"
IF R_RELAT = (1,5,6) AND != [ANY OTHER HH MEMBER] R_RELAT = 4 THEN
"ADULT"
IF R_RELAT = 8 AND [ANY OTHER HH MEMBER] (R_RELAT = (2,7) AND
IMPUTED_AGE > 21) THEN "CHILD"
IF R_RELAT = 8 AND != [ANY OTHER HH MEMBER] (R_RELAT = (2,7) AND
IMPUTED_AGE > 21) THEN "ADULT"
IF R_RELAT = (-7,-8) AND [ANY OTHER HH MEMBER] R_RELAT = 4 THEN "CHILD"
IF R_RELAT = (-7,-8) AND != [ANY OTHER HH MEMBER] R_RELAT = 4
IF R_RELAT = (2,7) THEN "ADULT"

HOUSEHOLD_ADULT_COUNT = COUNT (WHERE ADULT_CHILD_STATUS = "ADULT")
HOUSEHOLD_CHILD_COUNT = COUNT (WHERE ADULT_CHILD_STATUS = "CHILD")
HOUSEHOLD_RETIRED_COUNT = COUNT (WHERE PRMACT = 6 OR ((PRMACT IS
MISSING OR PRMACT = (-7,-8)) AND AGE >= 65))
MIN_AGE = MINIMUM (AGE)

LIF_CYC =
IF HOUSEHOLD_ADULT_COUNT = 1 AND HOUSEHOLD_CHILD_COUNT = 0 AND
HOUSEHOLD_RETIRED_COUNT = 0 THEN 1
IF HOUSEHOLD_ADULT_COUNT >= 2 AND HOUSEHOLD_CHILD_COUNT = 0 AND
HOUSEHOLD_RETIRED_COUNT = 0 THEN 2
IF HOUSEHOLD_ADULT_COUNT = 1 AND HOUSEHOLD_CHILD_COUNT >= 1 AND
MIN_AGE BETWEEN 0 AND 5 THEN 3
IF HOUSEHOLD_ADULT_COUNT >= 2 AND HOUSEHOLD_CHILD_COUNT >= 1 AND
MIN_AGE BETWEEN 0 AND 5 THEN 4
IF HOUSEHOLD_ADULT_COUNT = 1 AND HOUSEHOLD_CHILD_COUNT >= 1 AND
MIN_AGE BETWEEN 6 AND 15 THEN 5
IF HOUSEHOLD_ADULT_COUNT >= 2 AND HOUSEHOLD_CHILD_COUNT >= 1 AND
MIN_AGE BETWEEN 6 AND 15 THEN 6
IF HOUSEHOLD_ADULT_COUNT = 1 AND HOUSEHOLD_CHILD_COUNT >= 1 AND
MIN_AGE BETWEEN 16 AND 21 THEN 7
IF HOUSEHOLD_ADULT_COUNT >= 2 AND HOUSEHOLD_CHILD_COUNT >= 1 AND
MIN_AGE BETWEEN 16 AND 21 THEN 8
IF HOUSEHOLD_ADULT_COUNT = 1 AND HOUSEHOLD_CHILD_COUNT = 0 AND
HOUSEHOLD_RETIRED_COUNT = 1 THEN 9
IF HOUSEHOLD_ADULT_COUNT >= 2 AND HOUSEHOLD_CHILD_COUNT = 0 AND
HOUSEHOLD_RETIRED_COUNT >= 1 THEN 10

```

MSACAT

Description: Metropolitan Statistical Area (MSA) category for the household's home address, based on household's home geocode and TIGER/Line Shapefiles.

Logic: `IF MSASIZE = (4,5) THEN
IF RAIL = 1 THEN 1
IF RAIL = 2 THEN 2
IF MSASIZE = (1,2,3) THEN 3
IF CBSA IS MISSING THEN 4`

MSASIZE

Description: Population size category of the Metropolitan Statistical Area (MSA), from the 2010-2014 five-year American Community Survey (ACS) API.

Logic: `IF [POPULATION OF MSA] < 250,000 THEN "01"
IF [POPULATION OF MSA] >= 250,000 AND <= 499,999 THEN "02"
IF [POPULATION OF MSA] >= 500,000 AND <= 999,999 THEN "03"
IF [POPULATION OF MSA] >= 1,000,000 AND <= 2,999,999 THEN "04"
IF [POPULATION OF MSA] >= 3,000,000 THEN "05"
IF MSA IS MISSING THEN "06"`

NUMADLT

Description: Count of adult household members at least 18 years old

Logic: `[PERSON_IS_18_OVER] =
IF R_AGE >= 18 THEN TRUE
IF AGERANGE = (4,5,6) THEN TRUE

COUNT(WHERE PERSON_IS_18_OVER = TRUE)`

RAIL

Description: MSA heavy rail status for household

Logic: `IF CBSA =
("12060","12580","14460","39300","16980","17460","31080","37100","40140","33
100","35300","14860","37980","41860","41940","47900","35620") THEN "01"
ELSE "02"`

RESP_CNT

Description: Count of responding persons per household

Logic: HHSIZE

SMPLSRCE

Description: Sample where the case originated

Logic: `if SAMPAREA='0' then SMPLSRCE='01'; else SMPLSRCE='02';`

SCRESP

Description: Person identifier of mail screener respondent, always 1 to roster self first

Logic: 1

TDAYDATE

Description: Date of travel day (YYYYMM)

Logic: EXTRACT(YYYYMM FROM TDAYDAT2)

TRAVDAY

Description: Travel day - day of week

Logic: [DAY OF WEEK] = EXTRACT(DAY OF WEEK FROM TDAYDAT2)

```
IF [DAY OF WEEK] = SUNDAY THEN 1
IF [DAY OF WEEK] = MONDAY THEN 2
IF [DAY OF WEEK] = TUESDAY THEN 3
IF [DAY OF WEEK] = WEDNESDAY THEN 4
IF [DAY OF WEEK] = THURSDAY THEN 5
IF [DAY OF WEEK] = FRIDAY THEN 6
IF [DAY OF WEEK] = SATURDAY THEN 7
```

URBAN

Description: Household's urban area classification, based on home address and 2014 TIGER/Line Shapefile

Logic: IF [URBAN AREA TYPE] = "URBANIZED AREA" THEN "01"
IF [URBAN AREA TYPE] = "URBAN CLUSTER" THEN "02"
IF [GEOMETRY] SURROUNDED BY ([GEOMETRY] WHERE [URBAN AREA TYPE] = "URBANIZED AREA") THEN "03"
ELSE "04"

URBANSIZE

Description: Urban area size where home address is located

Logic: IF [POPULATION OF URBAN AREA] BETWEEN 50,000 AND 199,999 THEN 1
IF [POPULATION OF URBAN AREA] BETWEEN 200,000 AND 499,999 THEN 2
IF [POPULATION OF URBAN AREA] BETWEEN 500,000 AND 999,999 THEN 3
IF [POPULATION OF URBAN AREA] >= 1,000,000 THEN
IF RAIL = "01" THEN 4
IF RAIL = "02" THEN 5
ELSE 6

URBRUR

Description: Household in urban/rural area

Logic: IF URBAN = (01,02) THEN 1
ELSE 2

WEBUSE17

Description: Frequency of internet use

Logic: MINIMUM (
PC WHERE PC != (-7,-8),
SPHONE WHERE SPHONE != (-7,-8),
TAB WHERE TAB != (-7,-8),
ODEVICE WHERE ODEVICE != (-7,-8)
)

WRKCOUNT

Description: Number of workers in household

Logic: COUNT (WHERE WORKER = 1)

YOUNGCHILD

Description: Count of persons with an age between 0 and 4 in household

Logic: [PERSON_IS_0_TO_4] =
IF R_AGE = (0,1,2,3,4) THEN TRUE
IF AGERANGE = (1) THEN TRUE

COUNT(WHERE PERSON_IS_0_TO_4 = TRUE)

HTEEMPDN, HTHTNRNT, HTPPOPDN, HTRES DN, HBRES DN, HBPPOPDN, HBHTNRNT, HBHUR

Description: Tract and Block Group Variables (also known as Claritas Variables). These variables were added to describe the characteristics of the areas where the NHTS respondents were surveyed. This allows the data analyst to look for patterns in travel behavior, not only by individual characteristics, but also by neighborhood characteristics. The data user can examine how characteristics such as population density, housing density, renter occupancy rate, and urbanicity of the household location may affect individual travel behavior. These variables are referred to as the "Claritas" variables as the source of the data is Claritas, LLC. (Claritas). This is the third NHTS to have these Claritas data appended to enhance analysis, the first two being the 2001 and 2009 NHTS.

Logic: Sources of Tract and Block Group Variables: The data contained in these variables were derived from U.S. Decennial Census and American Community Survey data, enhanced by Claritas using data from regional and city planning agencies, federal agencies (e.g., Bureau of Labor Statistics, Bureau of Census, Bureau of Economic Analysis) U.S. Postal Service, the direct mail industry, the real estate industry, and experts in the fields of geographic information systems and mapmaking. These estimates are made at relatively small units of geography, such as census tracts and block groups, which make this update effective for use in supplementing the NHTS data based on the home location.

Variable Naming Scheme: The variable names were designed so that:

- many of these variables would fall together in an alphabetic listing, and
- the variable name would help in describing the contents.

The naming scheme is:

First letter	H for household descriptor
Second letter	B for block group level data
	T for tract level data
Third letter of Household variables	H for housing characteristic
	P for population characteristic
Last five letters	Describe the data in the variable, e.g. POPDN = population density; RESDN = residential density

For example, HTHRESDN is a household descriptor, at the tract level, describing a housing characteristic, specifically, residential density (RESDN).

The set of tract and block group variables derived by Claritas are:

Household Descriptor, Block Group Level

HBHRESDN	Housing units per square mile
HBHTNRNT	Percent renter-occupied housing
HBHUR	Urban/rural code (see below)
HBPPOPDN	Population density (persons per square mile)

Household Descriptor, Tract Level

HTEEMPDN	Employed persons per square mile (i.e., employed persons at their residence location)
HTHRESDN	Housing units per square mile
HTHTNRNT	Percent renter-occupied housing
HTPPOPDN	Population density (persons per square mile)

Urban-Rural Continuum: One of the most popular Claritas variables is a variable that describes a location based on where it falls on an urban-rural continuum that is categorized according to five levels of "urbanicity." This variable is HBHUR and represents urbanicity at the home location block group. This breakout of urban/rural should not be confused with the variable URBAN, which is the Census defined urban area status of the sample household.

The categories of the Claritas Urban/Rural Continuum and the distribution of NHTS households within these categories are presented in the following table, showing the distribution of the weighted and unweighted (number of respondent households) across the urban/rural continuum.

Urban/Rural Continuum – Distribution of Unweighted and Unweighted NHTS Sample

HBHUR	Unweighted NHTS Households	Percent of HHs unweighted	Weighted NHTS Households	Percent of households weighted
Second City	26,129	20.15	24,160,812	20.44
Rural	29,680	22.88	21,370,243	18.08
Suburban	28,824	22.22	26,249,815	22.21
Small Town	29,162	22.48	22,867,286	19.34
Urban	15,792	12.18	23,509,985	19.89
Not Ascertained	109	0.08	50,110	0.04
Total	129,696	100	118,208,251	100

For more information on urbanicity and its methodology please see https://nhts.ornl.gov/assets/Assessing_the_Role_of_Urbanicity.pdf .

PERSON Variables

ALT_16

Description: Alternative Mode of Transportation: Public Transportation or Taxi

Logic:

```
if ALT_1='01' and ALT_6='06' then ALT_16='03';
  else if ALT_1='01' then ALT_16='01';
  else if ALT_6='06' then ALT_16='02';
  else if ALT_1='-9' or ALT_6='-9' then ALT_16='-9';
    else ALT_16='04';
if r_age < 16 or proxy='02' then ALT_16='-1';
```

ALT_23

Description: Alternative Mode of Transportation: Passenger to Friend/Family Member or Rental Car

Logic:

```
if ALT_2='02' and ALT_3='03' then ALT_23='03';
  else if ALT_2='02' then ALT_23='01';
  else if ALT_3='03' then ALT_23='02';
  else if ALT_2='-9' or ALT_3='-9' then ALT_23='-9';
    else ALT_23='04';
if r_age < 16 or proxy='02' then ALT_23='-1';
```

ALT_45

Description: Alternative Mode of Transportation: Bicycle or Walk

Logic:

```
if ALT_4='04' and ALT_5='05' then ALT_45='03';
  else if ALT_4='04' then ALT_45='01';
  else if ALT_5='05' then ALT_45='02';
  else if ALT_4='-9' or ALT_5='-9' then ALT_45='-9';
    else ALT_45='04';
if r_age < 16 or proxy='02' then ALT_45='-1';
```

BIKE_DFR

Description: Reason for Not Biking More: Infrastructure

Logic:

```
if BIKE_D='01' and BIKE_R='02' and BIKE_F='03' then
  BIKE_DFR='07';
  else if BIKE_D='01' and BIKE_F='03' then BIKE_DFR='04';
  else if BIKE_D='01' and BIKE_R='02' then BIKE_DFR='05';
  else if BIKE_R='02' and BIKE_F='03' then BIKE_DFR='06';
  else if BIKE_D='01' then BIKE_DFR='01';
  else if BIKE_F='03' then BIKE_DFR='02';
  else if BIKE_R='02' then BIKE_DFR='03';
  else BIKE_DFR='08';
if BIKE_D='-1' and BIKE_R='-1' and BIKE_F='-1' then BIKE_DFR='-1';
if BIKE_D='-9' and BIKE_R='-9' and BIKE_F='-9' then BIKE_DFR='-9';
```

BIKE_GKP

Description: Reason for Not Biking More: Safety

Logic:

```
if BIKE_G='04' and BIKE_K='05' and BIKE_P='06' then
    BIKE_GKP='07';
else if BIKE_G='04' and BIKE_K='05' then BIKE_GKP='04';
else if BIKE_G='04' and BIKE_P='06' then BIKE_GKP='05';
else if BIKE_K='05' and BIKE_P='06' then BIKE_GKP='06';
else if BIKE_G='04' then BIKE_GKP='01';
else if BIKE_K='05' then BIKE_GKP='02';
else if BIKE_P='06' then BIKE_GKP='03';
else BIKE_GKP='08';
if BIKE_G='-1' and BIKE_K='-1' and BIKE_P='-1' then BIKE_GKP='-1';
if BIKE_G='-9' and BIKE_K='-9' and BIKE_P='-9' then BIKE_GKP='-9';
```

CNTTDTR

Description: Count of person trips on travel day

Logic: COUNT(TDTRPNUM)

DIARY

Description: Travel Diary completion status

Logic: IF DIARYHAV = 1 AND DIARYCMP = 1 THEN 1
ELSE 2

DRIVER

Description: Driver status, derived

Logic: IF R_AGE > 0 AND R_AGE < 15 THEN 2
IF AGERANGE = (1,2) THEN 2
IF DRVR = 1 THEN 1
IF COUNT(TDTRPNUM WHERE WHODROVE = PERSONID) > 0 THEN 1
IF COUNT(TDTRPNUM WHERE WHODROVE = PERSONID) = 0 THEN 2

FRSTHM17

Description: Travel day began at home location

Logic: IF LOCATION IS HOME WHERE TDTRPNUM = 1 THEN 1
IF LOCATION IS NOT HOME WHERE TDTRPNUM = 1 THEN 2

GCDWORK

Description: Minimum geodesic (Great Circle) distance between home location and work location in meters, using WGS84 coordinate system

Logic: [HOME_LOCATION] = (LONGITUDE, LATITUDE) WHERE LOCATION.LOCTYPE = 1
[WORK_LOCATION] = (LONGITUDE, LATITUDE) WHERE PERSON.PERSONID = LOCATION.PERSONID AND LOCATION.LOCTYPE = 2
https://geographiclib.sourceforge.io/html/C/geodesic_8h.html#a19bc3d000428010ad9d8509174e672c9

OUTOFTWN

Description: Away from home for entire travel day

Logic: `IF COUNT(TDTRPNUM WHERE LOCATION IS HOME) = 0 THEN 1
ELSE 2`

R_RACE

Description: Race

Logic: `IF COUNT (RACE_*) > 1 THEN 6
IF COUNT (RACE_*) = 1 THEN RACE`

USEPUBTR

Description: Public Transit Usage on Travel Date, derived

Logic: `IF COUNT(TRPTRANS17) WHERE TRPTRANS17 = (11,15,16) > 0 THEN 1
ELSE 2`

WALK_DEF

Description: Reason for Not Walking More: Infrastructure

Logic: `if WALK_D='01' and WALK_E='02' and WALK_F='03' then
 WALK_DEF='07';
 else if WALK_D='01' and WALK_E='02' then WALK_DEF='04';
 else if WALK_D='01' and WALK_F='03' then WALK_DEF='05';
 else if WALK_E='02' and WALK_F='03' then WALK_DEF='06';
 else if WALK_D='01' then WALK_DEF='01';
 else if WALK_E='02' then WALK_DEF='02';
 else if WALK_F='03' then WALK_DEF='03';
 else WALK_DEF='08';
if WALK_D='-1' and WALK_E='-1' and WALK_F='-1' then WALK_DEF='-
1';
if WALK_D='-9' and WALK_E='-9' and WALK_F='-9' then WALK_DEF='-
9';`

WALK_GKQ

Description: Reason for Not Walking More: Safety

Logic: `if WALK_G='04' and WALK_K='05' and WALK_Q='06' then
 WALK_GKQ='07';
 else if WALK_G='04' and WALK_K='05' then WALK_GKQ='04';
 else if WALK_G='04' and WALK_Q='06' then WALK_GKQ='05';
 else if WALK_K='05' and WALK_Q='06' then WALK_GKQ='06';
 else if WALK_G='04' then WALK_GKQ='01';
 else if WALK_K='05' then WALK_GKQ='02';
 else if WALK_Q='06' then WALK_GKQ='03';
 else WALK_GKQ='08';
if WALK_G='-1' and WALK_K='-1' and WALK_Q='-1' then WALK_GKQ='-
1';
if WALK_G='-9' and WALK_K='-9' and WALK_Q='-9' then WALK_GKQ='-
9';`

WKSTFIPS

Description: The state FIPS code for the respondent's geocoded work address. The state FIPS codes were identified using United States Census Bureau 2016 TIGER/Line Shapefiles.

Logic: NA

WORKER

Description: Worker status

Logic:
IF (R_AGE >= 0 AND < 16 OR AGERANGE = (1,2) THEN -1
IF (PRMACT = (-8,-7) AND PAYPROF = (-8,-7)) OR (PRMACT IS NULL AND (R_AGE > 15 OR AGERANGE = (3,4,5,6))) THEN -9
IF PRMACT = (1,2) OR PAYPROF = 1 THEN 1
ELSE 2

R_AGE_IMP

Description: Age (imputed)

Logic: Age of subject used in weighting. Replace values in R_AGE that are -7, -8 or -9 with the imputed age values.

R_SEX_IMP

Description: Gender (imputed)

Logic: Gender of subject used in weighting. Replace values in R_SEX that are -7, -8 or -9 with the imputed gender values.

TRIP Variables

DBHTNRNT, DBHUR, DBPPOPDN, DBRES DN, DTEEMPDN, DTHTNRNT, DTPPOPDN, and DTRES DN

Description: Density and demographic variables based on the Tract and Block Group of the **trip destination**. These are in the same format as those found for the household location (see description in the Household file for HTEEMPDN, HTHTNRNT, HTPPOPDN, HTRES DN, HBRES DN, HBPPOPDN, HBHTNRNT, HBHUR).

DRVR_FLG

Description: Respondent drove on trip

Logic: IF PERSONID = WHODROVE AND TRPTRANS17 = (3,4,5,6,7,8,9,18) THEN 1
IF PERSONID != WHODROVE AND TRPTRANS17 = (3,4,5,6,7,8,9,18) THEN 2
IF TRPTRANS17 != (3,4,5,6,7,8,9,18) THEN -1

DWELTIME

Description: Time at destination

Logic: DEPTIME - ENDTIME

ENDTIME

Description: Trip End Time (HHMM)

Logic: Trip end time in military format.

GASPRICE

Description: Weekly regional gasoline price during the week of the household's travel day. Price of gasoline, in cents, on respondent's travel day.

Logic: PADD_REGION =
IF HHSTATE = ("CT","ME","MA","NH","RI","VT") THEN "PADD1A"
IF HHSTATE = ("DE","DC","MD","NJ","NY","PA") THEN "PADD1B"
IF HHSTATE = ("FL","GA","NC","SC","VA","WV") THEN "PADD1C"
IF HHSTATE =
("IL","IN","IA","KS","KY","MI","MN","MO","NE","ND","SD","OH","OK","TN","WI")
THEN "PADD2"
IF HHSTATE = ("AL","AR","LA","MS","NM","TX") THEN "PADD3"
IF HHSTATE = ("CO","ID","MT","UT","WY") THEN "PADD4"
IF HHSTATE = ("AK","AZ","CA","HI","NV","OR","WA") THEN "PADD5"
Source: EIA database (<https://www.eia.gov/petroleum/gasdiesel/>)

HHMEMDRV

Description: Household member drove on trip

Logic: IF TRPTRANS17 != (3,4,5,6,7,8,9,18) THEN -1
IF WHODROVE != 97 AND TRPTRANS17 = (3,4,5,6,7,8,9,18) THEN 1
IF WHODROVE = 97 AND TRPTRANS17 = (3,4,5,6,7,8,9,18) THEN 2

HH_ONTD

Description: Number of household members on trip including respondent

Logic: TRPHHACC + 1

LOOP_TRIP

Description: Trip origin and destination at identical location

Logic: IF TRIP DOES NOT BEGIN AND END AT THE SAME LOCATION THEN 1
ELSE 2

NONHHCNT

Description: Number of non-household members on trip

Logic: TRPACCM - TRPHHACC

NUMONTRP

Description: Number of people on trip including respondent

Logic: TRPACCM + 1

OBHTNRNT, OBHUR, OBPPOPDN, OBRES DN, OTEEMP DN, OTHTNRNT, OTPPOPDN, and OTRES DN

Description: Density and demographic variables based on the Tract and Block Group of the **trip origin**. These are in the same format as those found for the household location (see description in the Household file for HTEEMP DN, HTHTNRNT, HTPPOP DN, HTRES DN, HBRES DN, HBPPOP DN, HBHTNRNT, HBHUR).

PSGR_FLG

Description: Respondent was passenger on trip

Logic: IF WHODROVE != PERSONID AND TRPTRANS17 = (3,4,5,6,7,8,9,18) THEN 1
IF WHODROVE = PERSONID AND TRPTRANS17 = (3,4,5,6,7,8,9,18) THEN 2
IF TRPTRANS17 != (3,4,5,6,7,8,9,18) THEN -1

PUBTRANS

Description: Public transportation used on trip

Logic: IF TRPTRANS17 = (11,15,16) THEN 1
ELSE 2

STRTIME

Description: Trip Start Time (HHMM)

Logic: Trip start time in military format

TDTRPNUM

Description: Incrementing travel day trip number, starting at 1 for each person in the file

Logic: FOR EACH (HOUSEID, PERSONID) ORDERED BY STRTTIME (ROW NUMBER)

TDWKND

Description: Weekend trip

Logic: IF [TRAVDAY] = (1,7) OR ([TRAVDAY] = (6) AND STRTTIME >= 1800) THEN 1
ELSE 2

TRACC_WLK

Description: Walk as mode used to get to public transit?

Logic: if trprtrans in ('11', '15', '16') then do;
if TRACC1='01' then TRACC_WLK='01';
else if TRACC1='-9' then TRACC_WLK='-9';
else TRACC_WLK='02';
end;
else TRACC_WLK='-1';

TRACC_POV

Description: POV as mode used to get to public transit?

Logic: if trprtrans in ('11', '15', '16') then do;
if (TRACC3='03' or TRACC4='04' or TRACC5='05' or TRACC6='06' or TRACC7='07' or TRACC8='08' or TRACC9='09' or TRACC18='18') then TRACC_POV='01';
else if TRACC3='-9' then TRACC_POV='-9';
else TRACC_POV='02';
end;
else TRACC_POV='-1';

TRACC_BUS

Description: Bus as mode used to get to public transit?

Logic: if trprtrans in ('11', '15', '16') then do;
if TRACC11='11' then TRACC_BUS='01';
else if TRACC11='-9' then TRACC_BUS='-9';
else TRACC_BUS='02';
end;
else TRACC_BUS='-1';

TRACC_CRL

Description: Rail as mode used to get to public transit?

Logic: if trprtrans in ('11', '15', '16') then do;
if TRACC15='15' then TRACC_CRL='01';
else if TRACC15='-9' then TRACC_CRL='-9';
else TRACC_CRL='02';
end;
else TRACC_CRL='-1';

TRACC_SUB

Description: Subway as mode used to get to public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if TRACC16='16' then TRACC_SUB='01';
    else if TRACC16='-9' then TRACC_SUB='-9';
    else TRACC_SUB='02';
end;
else TRACC_SUB='-1';
```

TRACC_OTH

Description: Other mode used to get to public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if (TRACC2='02' or TRACC10='10' or TRACC12='12' or
    TRACC13='13' or TRACC14='14' or TRACC17='17' or TRACC18='18' or
    TRACC19='19' or TRACC20='20' or TRACC_O='97') then
    TRACC_OTH='01';
    else if TRACC2='-9' then TRACC_OTH='-9';
    else TRACC_OTH='02';
end;
else TRACC_SUB='-1';
```

TREGR_WLK

Description: Walk as mode used to get from public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if TREGR1='01' then TREGR_WLK='01';
    else if TREGR1='-9' then TREGR_WLK='-9';
    else TREGR_WLK='02';
end;
else TREGR_WLK='-1';
```

TREGR_POV

Description: POV as mode used to get from public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if (TREGR3='03' or TREGR4='04' or TREGR5='05' or
    TREGR6='06' or TREGR7='07' or TREGR8='08' or TREGR9='09' or
    TREGR18='18') then TREGR_POV='01';
    else if TREGR3='-9' then TREGR_POV='-9';
    else TREGR_POV='02';
end;
else TREGR_POV='-1';
```

TREGR_BUS

Description: Bus as mode used to get from public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if TREGR11='11' then TREGR_BUS='01';
    else if TREGR11='-9' then TREGR_BUS='-9';
    else TREGR_BUS='02';
end;
else TREGR_BUS='-1';
```

TREGR_CRL

Description: Rail as mode used to get from public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if TREGR15='15' then TREGR_CRL='01';
    else if TREGR15='-9' then TREGR_CRL='-9';
    else TREGR_CRL='02';
end;
else TREGR_CRL='-1';
```

TREGR_SUB

Description: Subway as mode used to get from public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if TREGR16='16' then TREGR_SUB='01';
    else if TREGR16='-9' then TREGR_SUB='-9';
    else TREGR_SUB='02';
end;
else TREGR_SUB='-1';
```

TREGR_OTH

Description: Other mode used to get from public transit?

Logic:

```
if trprtrans in ('11','15','16') then do;
    if (TREGR2='02' or TREGR10='10' or TREGR12='12' or
    TREGR13='13' or TREGR14='14' or TREGR17='17' or TREGR18='18' or
    TREGR19='19' or TREGR20='20' or TREGR_O='97') then
    TREGR_OTH='01';
    else if TREGR2='-9' then TREGR_OTH='-9';
    else TREGR_OTH='02';
end;
else TREGR_OTH='-1';
```

TRIPPURP

Description: Generalized purpose of trip, home-based and non-home based

Logic:

```
IF WHYFROM = -9 OR WHYTO = -9 THEN -9
IF WHYFROM = (1,2) AND WHYTO = (3,4) THEN HBW
IF WHYFROM = (3,4) AND WHYTO = (1,2) THEN HBW
IF WHYFROM = (1,2) AND WHYTO = (11,12,13) THEN HBSHP
IF WHYFROM = (11,12,13) AND WHYTO = (1,2) THEN HBSHP
IF WHYFROM = (1,2) AND WHYTO = (15,16,17) THEN HBSOC
IF WHYFROM = (15,16,17) AND WHYTO = (1,2) THEN HBSOC
IF WHYFROM = (1,2) AND WHYTO != (3,4,11,12,13,15,16,17) THEN HBO
IF WHYFROM != (3,4,11,12,13,15,16,17) AND WHYTO = (1,2) THEN HBO
ELSE NHB
```

TRPHHACC

Description: Count of Household Members on Trip

Logic: COUNT(ONTD_P* = 1)

TRPMILAD

Description: Trip distance in miles, adjusted for comparability to past surveys

Logic: `IF TRPTRANS IN (3,4,5,6,8,9,18) THEN
IF WHYTRP90 = 01 THEN TRPMILES * 1.0735
IF WHYTRP90 != 01 THEN TRPMILES * 1.1151
ELSE TRPMILES`

TRPMILES

Description: Trip distance in miles, derived from route geometry returned by Google Maps API, or from reported loop-trip distance

Logic: `IF TRIP DOES NOT BEGIN AND END AT THE SAME LOCATION THEN
[GOOGLE_ROUTE_DISTANCE]
IF TRPTRANS17 IN (1,2,3,4,5,6,7,8,9,18) AND WKBK_UNIT = (1,2) THEN
IF WKBK_UNIT = 1 THEN WKBK_DIST * 0.1111
IF WKBK_UNIT = 2 THEN WKBK_DIST
ELSE [GOOGLE_ROUTE_DISTANCE]`

TRPTRANS

Description: Trip Mode, derived

Logic: `[VEHICLE TYPE] = VEHICLE.VEHTYPE WHERE VEHICLE.VEHID = TRIP.VEHID

IF [VEHICLE TYPE] = 1 THEN 3
IF [VEHICLE TYPE] = 2 THEN 5
IF [VEHICLE TYPE] = 3 THEN 4
IF [VEHICLE TYPE] = 4 THEN 6
IF [VEHICLE TYPE] = 5 THEN 6
IF [VEHICLE TYPE] = 6 THEN 9
IF [VEHICLE TYPE] = 7 THEN 8
ELSE TRPTRANS17`

TRVLCMIN

Description: Trip Duration in Minutes

Logic: `[WKBK_DISTANCE] =
IF WKBK_UNIT = 2 THEN WKBK_DIST
IF WKBK_UNIT = 1 THEN WKBK_DIST * 0.111111`

Derived from STRTTIME and ENDTIME

VMT_MILE

Description: Trip distance in miles for personally driven vehicle trips, derived from route geometry returned by Google Maps API

Logic: `IF TRPTRANS17 = (3,4,5,6,7,8,9,18) AND VEHTYPE = (1,2,3,4,5,6,7) AND
DRVR_FLG = 1 THEN TRPMILES
ELSE -1`

WHYFROM

Description: Trip Origin Purpose

Logic: `WHYTO WHERE TDTRPNUM = TDTRPNUM - 1`

WHYTRP1S

Description: Trip purpose summary

Logic:

```
IF WHYTO = (1,2) THEN 01
IF WHYTO = (3,4) THEN 10
IF WHYTO = (8,9,10,19) THEN 20
IF WHYTO = 18 THEN 30
IF WHYTO = (11,12,14) THEN 40
IF WHYTO = (15,16,17) THEN 50
IF WHYTO = 6 THEN 70
IF WHYTO = 13 THEN 80
ELSE 97
```

WHYTRP90

Description: Travel day trip purpose consistent with 1990 NPTS design.

Logic:

```
[HOME TOUR WINDOW] = MINIMUM(TDTRPNUM) WHERE WHYTO IN (1, 2) BETWEEN
MAXIMUM(TDTRPNUM) WHERE WHYTO IN (1, 2)
[WORK TOUR WINDOW] = MINIMUM(TDTRPNUM) WHERE WHYTO IN (3) BETWEEN
MAXIMUM(TDTRPNUM) WHERE WHYTO IN (3)
```

```
[WHYTRP90 STEP ONE] =
IF WHYTO IN (1,2) AND [HOME TOUR WINDOW] = 0 THEN O_WHYTO
IF WHYTO IN (1,2) AND [HOME TOUR WINDOW] >= 1 THEN WHYTO WHERE
DWELLTIME = MAXIMUM(DWELLTIME) OF [HOME TOUR WINDOW]
IF WHYTO IN (3) AND [WORK TOUR WINDOW] = 0 THEN WHYTO
IF WHYTO IN (3) AND [WORK TOUR WINDOW] >= 1 THEN WHYTO WHERE
DWELLTIME = MAXIMUM(DWELLTIME) OF [WORK TOUR WINDOW]
IF WHYTO IN (1,2) AND WHYFROM IN (1,2) AND LOOP_TRIP = 1 AND TRPTRANS IN
(1,2) THEN 16
ELSE WHYTO
```

```
WHYTRP90 =
IF [WHYTRP90 STEP ONE] = (3) THEN 1
IF [WHYTRP90 STEP ONE] = (4) THEN 2
IF [WHYTRP90 STEP ONE] = (11) THEN 3
IF [WHYTRP90 STEP ONE] = (5,6,10,12,14) THEN 4
IF [WHYTRP90 STEP ONE] = (13) AND IN [WORK TOUR WINDOW] THEN 4
IF [WHYTRP90 STEP ONE] = (8,9,19) THEN 5
IF [WHYTRP90 STEP ONE] = (18) THEN 6
IF [WHYTRP90 STEP ONE] = (17) THEN 8
IF [WHYTRP90 STEP ONE] = (15,16) THEN 10
IF [WHYTRP90 STEP ONE] = (13) AND IN [HOME TOUR WINDOW OR NOT WORK
TOUR WINDOW] THEN 10
IF [WHYTRP90 STEP ONE] = (97) THEN 11
IF [WHYTRP90 STEP ONE] = (-8,-7) THEN 99
ELSE 11
```

VEHICLE Variables

ANNMILES

Description: Self-reported annualized mile estimate

Logic: IF VEHOWNED = (-7,-8) OR VEHMILES = (-7,-8,-77,-88) OR ESTMILES = (-7,-8,-77,-88) THEN -9
IF VEHOWNED = 1 THEN VEHMILES
IF VEHOWNED = 2 THEN ROUND((12 * ESTMILES) / OWNUNIT)

BESTMILE

Description: Best estimate of annual miles

Logic: See <https://nhts.ornl.gov/documentation> for information on the 2017 Best Estimate of Annual Vehicle Mileage (BESTMILE) and its associated flags

BEST_EDT

Description: Flag any edits/adjustments to BESTMILE

Logic: See <https://nhts.ornl.gov/documentation> for information on the 2017 Best Estimate of Annual Vehicle Mileage (BESTMILE) and its associated flags

BEST_FLG

Description: How BESTMILE was computed

Logic: See <https://nhts.ornl.gov/documentation> for information on the 2017 Best Estimate of Annual Vehicle Mileage (BESTMILE) and its associated flags

BEST_OUT

Description: Flag identifying BESTMILE outlier values

Logic: See <https://nhts.ornl.gov/documentation> for information on the 2017 Best Estimate of Annual Vehicle Mileage (BESTMILE) and its associated flags

FEGEMPG

Description: Fuel Economy.gov EIA-Derived 55/45 fuel economy

Logic: See <https://nhts.ornl.gov/documentation> for information on 2017 EIA Fuel Economy and Fuel Use variables

FEGEMPGA

Description: Fuel Economy.gov 55/45 alternative fuel economy

Logic: See <https://nhts.ornl.gov/documentation> for information on 2017 EIA Fuel Economy and Fuel Use variables

FEGEMPGF

Description: Flag for FEGEMPG

Logic: See <https://nhts.ornl.gov/documentation> for information on 2017 EIA Fuel Economy and Fuel Use variables

GSCOST

Description: Annualized fuel cost in US dollars per equivalent gallon

Logic: See <https://nhts.ornl.gov/documentation> for information on 2017 EIA Fuel Economy and Fuel Use variables

GSTOTCST

Description: Annual fuel expenditures in US dollars

Logic: See <https://nhts.ornl.gov/documentation> for information on 2017 EIA Fuel Economy and Fuel Use variables

GSYRGAL

Description: Annual fuel consumption in US gallons

Logic: See <https://nhts.ornl.gov/documentation> for information on 2017 EIA Fuel Economy and Fuel Use variables

HYBRID

Description: Hybrid vehicle

Logic: IF FUEL = 3 THEN 1
IF FUEL != 3 THEN 2

VEHAGE

Description: Age of vehicle, based on model year

Logic: IF VEHYEAR = (-8,-7) THEN VEHYEAR
IF VEHYEAR = (2016,2017,2018) THEN 1
ELSE 2017-VEHYEAR